

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.
OM protein - protein search, using sw model
Run On: February 11, 2003, 19:44:29 ; Search time 28.6571 Seconds
(without alignments)
3178.015 Million cell updates/sec
Title: US-09-497-967-6
Perfect score: 2342
Sequence: 1 MKYNILLIILISLFINELRA.....STTFAKFLISLLIFSYLL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2342	100.0	442	5	Q9XZG2
2	2087	89.1	395	5	Q27208
3	1333	56.9	460	5	Q962N5
4	921	39.3	468	5	Q9BMH3
5	281.5	12.0	371	5	Q9GPP0
6	231	9.9	316	5	Q9GPP3
7	230	9.8	316	5	Q9GPP4
8	229	9.8	305	5	Q9GPP2
9	222	9.5	560	5	Q9U013
10	219.5	9.4	548	5	Q9GQ45
11	214.5	9.2	709	5	Q9XTJ7
12	212.5	9.1	704	5	Q9U048
13	212	9.1	1274	5	Q24977
14	210.5	9.0	557	5	Q24992
15	210.5	9.0	597	5	Q07317
16	210.5	9.0	1154	5	Q9GQ46

17	210.5	9.0	1274	5	Q9NGL3
18	206	8.8	594	5	Q24970
19	205.5	8.8	709	5	Q97444
20	205	8.8	1459	5	Q17084
21	199	8.5	397	5	Q95V71
22	198.5	8.5	423	5	Q9U697
23	196.5	8.4	719	5	Q9U021
24	195.5	8.3	1372	5	P91526
25	194.5	8.3	350	5	Q94589
26	194.5	8.3	667	5	Q9XTK3
27	194.5	8.3	719	5	Q9U019
28	194.5	8.3	1168	5	Q96758
29	191.5	8.2	436	5	Q27197
30	191.5	8.2	1551	5	Q9NGV4
31	191.5	8.2	3396	5	Q9VM55
32	190	8.1	5374	11	Q99ND0
33	189	8.1	645	5	Q97448
34	187.5	8.0	421	5	Q95V69
35	186.5	8.0	504	5	Q95V69
36	184	7.9	394	5	Q9GQ47
37	183	7.8	769	5	Q24971
38	183	7.8	1299	5	Q28489
39	182	7.7	1622	5	Q06550
40	180.5	7.7	2395	5	Q27167
41	179	7.6	3695	4	Q8TDF8
42	177	7.6	439	5	Q94827
43	176	7.5	419	5	Q8SYE6
44	176	7.5	448	5	Q9W491
45	176	7.5	503	5	Q9U018

ALIGNMENTS

RESULT 1

Q9XZG2	ID	Q9XZG2	PRELIMINARY:	PRT;	442 AA.
AC	Q9XZG2;				
DT	01-NOV-1999 (Tremblrel. 12, Created)				
DT	01-NOV-1999 (Tremblrel. 12, Last sequence update)				
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)				
DE	Immunization antigen precursor.				
GN	IAG48.				
OS	Ichthyophthirius multifiliis.				
OC	Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;				
OC	Ophryoglenina; Ichthyophthirius.				
OX	NCBI_TaxID=5932;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-G1;				
RX	MEDLINE=99196987; PubMed=10095108;				
RA	Clark T.G., Lin T.L., Jackwood D.A., Sherrill J., Lin Y.,				
RA	Dickerson H.W.;				
RT	"The gene for an abundant parasite coat protein predicts tandemly				
RT	repetitive metal binding domains.";				
RL	Gene 229:91-100(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-G1;				
RA	Gaertig J., Gao Y., Tishgarten T., Clark T.G., Dickerson H.W.;				
RT	"Surface display of a parasite antigen in the ciliate Tetrahymena				
RT	thermophila.";				
RL	Nat. Biotechnol. 0:0-0(1999).				
DR	EMBL: AF140273; AA031283.1; -.				
DR	PRINTS: PR01574; TUBBYPROTEIN.				
KW	SIGNAL.				
FT	CHAIN	1	20	POTENTIAL.	
FT	CHAIN	21	442	IMMOBILIZATION ANTIGEN	
SQ	SEQUENCE	442 AA;	45025 MW;	52658FP65D27AFA CRC64;	

Query Match

Best Local Similarity 100.0%; Score 2342; DB 5; Length 442;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

RESULT 6
Q9GPP3 PRELIMINARY; PRT; 316 AA.
AC Q9GPP3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Immobilization antigen LB (Fragment).
GN SERLB.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP STRAIN=B;
RC MEDLINE=20549003; PubMed=11095959;
RA Doerder F.P.; Gerber C.A.;
RT "Molecular Characterization of the SerL Paralogs of Tetrahymena
thermophila."
RL Biochem. Biophys. Res. Commun. 278:621-626(2000).
DR EMBL; AF312771; AAG38117.1; -.
FT NON_TER
SQ SEQUENCE 316 AA; 3008 MW; 133A0B7D0797A3BD CRC64;

Query Match 9.9%; Score 231; DB 5; Length 316;
Best Local Similarity 27.7%; Pred. No. 9.1e-09;
Matches 106; Conservative 24; Mismatches 127; Indels 126; Gaps 21;

QY 7 LILIIISLFTNEIRAVPCPDGTQAGLDVGAADLGTCVNCPRNFYNGGAQGEANGNQ 66
Db 6 LILISLAVIATVNA--CTD-----TNATAGAGGTCTF--CNAGY-----GTSTDVT 47
QY 67 PFAANNAARGICVPCQINRVGSVTNAGDLATLATQCSQCTGTALDDGVDVDFDSAAQ 126
Db 48 P-----SGSCTKCTPTGNTSAAATAS--GTLVSSC-----TCNDTNASLKGDNSSGCQ 91
QY 127 CVKCKPNFYNGSGPGEAPGVQVFAAGAAAGVAAVTSQCVPCQLNKNDSPATAGAAQ 186
Db 92 ---CKANFY---GTPNAVSGG-----ATGCG-TACPTGSAAGST 179
QY 187 LATQCSNQCTGTALDDGVTLVF-----NTSATL-----CVKCRPNFYNGSGPGEAP 235
Db 107 -ATGCS-ACPTGTTSPAGTAATVSCACNDTNASLKGDNSSGC-QCKANFY---GTPNAVSG 163
QY 236 GVQVFAAG-AAAAGVAAVTSQCVPCQINKNDSP-ATGCG-TACPTGSAAGST 279
Db 164 GCTACPTGSAAGAAAGSTAVTS---CACNDTNASLKGDNSSGC-QCKANFY---GTPNAVSG 219
QY 280 CSTQCTGTALDDGVTLVF-----NST-OCSCIANFYFNGNFEAGKSCQCLKC 328
Db 220 C-TACPTGTTSTAGTVIGSCACPDNTAALNSATPVCCOKAN-FYGIPTASGASGCTAC 277
QY 329 PVSKTTPAHAPGNATQATQCLT 351
Db 278 PSGOTAPAGSATNVCKAASTST 300

RESULT 7
Q9GPP4 PRELIMINARY; PRT; 316 AA.
AC Q9GPP4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Immobilization antigen LA (Fragment).
GN SERLA.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=20549003; PubMed=11095959;
RA Doerder F.P.; Gerber C.A.;
RT "Molecular Characterization of the SerL Paralogs of Tetrahymena
thermophila."
RL Biochem. Biophys. Res. Commun. 278:621-626(2000).
DR EMBL; AF312770; AAG38116.1; -.
FT NON_TER
SQ SEQUENCE 316 AA; 30100 MW; 1A13D076F28ED3BD CRC64;

Query Match 9.8%; Score 230; DB 5; Length 316;
Best Local Similarity 25.3%; Pred. No. 1.1e-08;
Matches 116; Conservative 34; Mismatches 138; Indels 170; Gaps 27;

QY 7 LILIIISLFTNEIRAVPCPDGTQAGLDVGAADLGTCVNCPRNFYNGGAQGEANGNQ 66
Db 6 LILISLAVIATVNA--CTD-----TNATAGAGGTCTF--CNAGY-----GTSTDVT 47
QY 67 PFAANNAARGICVPCQINRVGSVTNAGDLATLATQCSQCTGTALDDGVDVDFDSAAQ 126
Db 48 P-----SGSCTKCTPTGNTSAAATAS--GTLVSSC-----TCNDTNASLKGDNSSGCQ 91
QY 127 CVKCKPNFYNGSGPGEAPGVQVFAAGAAAGVAAVTSQCVPCQLNKNDSPATAGAAQ 186
Db 92 ---CKANFY---GTPNAVSGG-----ATGCG-TACPTGSAAGST 179
QY 236 GVQVFAAGAAAGVAAVTSQCVPCQINKNDSPATAGAAQANLATQCSQCTGTALDDGVT 295
Db 161 G-----ATGCG-TACPTGSAAGST 179
QY 296 LVFS-----NSTQCSQ---CIANYFFNGNFEA-GKSOCLKCPVSKTTPAHAPGNAT 344
Db 180 AVTSCACNDTNSSLKADNSACVCKANFYGTGTPNAVAGATGCTACTP-----GTST 230
QY 345 QATQCLTTPAGTVLDDGTSTNFVASATECTKCSAGFPASKTGTTAGTGTCTCTCKLT 404
Db 231 AGTTVIGSCACP---DTNASLN-TATPPVC-QCNANFYGTPT---SSGASGCTAC---P 278
QY 405 SGATAKYAEATQKVCQASTTFEAKLSLFLSFYLL 442
Db 279 SGTAPAPA-GSATNVCKAASTSTYIILPVSLLFLSLVML 315

RESULT 8
Q9GPP2 PRELIMINARY; PRT; 305 AA.
AC Q9GPP2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Immobilization antigen LC (Fragment).
GN SERLC.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=20549003; PubMed=11095959;
RA Doerder F.P.; Gerber C.A.;
RT "Molecular Characterization of the SerL Paralogs of Tetrahymena
thermophila."
RL Biochem. Biophys. Res. Commun. 278:621-626(2000).
DR EMBL; AF312772; AAG38118.1; -.
FT NON_TER
SQ SEQUENCE 305 AA; 28863 MW; 0568C353A0253564 CRC64;

```

```
Query Match          9.8%; Score 229; DB 5; Length 305;
Best Local Similarity 27.7%; Pred. No. 1.2e-08;
Matches 106; Conservative 24; Mismatches 127; Indels 126; Gaps 20;

QY 7 LILISLFINELRAVPCDGTQTQAGLTDYGAADLGTVCNCRPNFYNGGAAQGEANGQ 66
Db 6 LILISLAVIATVNA--CTD-----TNATAGAGTCTCF--CNAGYY-----GTSTDYV 47
QY 67 PFAANNAARGTCVPCQINRVGSVTNAGDLATLATATQSTQCTPTGALDDGVTDFDRSAAQ 126
Db 48 P-----SGCTKCP-----TGTSVAATASTGLVSSCTCNDTNASLKGDNSSGQ 91
QY 127 CVKCKPNFYNGSGPOGEAPGVQVFAAGAAAGVAAVTSQVPCQLNKNDSPATAGAAQAN 186
Db 92 ---CKANFY---GTPNAVSGG-----NTSATL-----CVKCRPNFYNGSGPOGEAP 235
QY 187 LATCSNOCPTGTVDLGGVTLVF-----NTSATL-----CVKCRPNFYNGSGPOGEAP 235
Db 107 -ATGCS-ACPTGTTSPAGTAAVTSACACNDTNASLKGDNSSGC-QCKANFYGTTPNAVAGGAT 163
QY 236 GVQVFAAG-AAAAGVAAVTSQVPCQINKNDS-----ATAGAAQANLATQ 279
Db 164 GCTACPTGSAAGAASTAVTS-----CACNDTNSLKAACVCKANFYGTTPNAVAGGATG 219
QY 280 CSTOCPGTATQDGVTLVFS-----NSST-OCSCICIANFYFNGFNFEAGKSQLCKC 328
Db 220 C-TACPTGTTSTAGTTVIGSCACPDNTAALNSATPPVQCCKAN-FVGTPTASGSGCTAC 277
QY 329 PVSKTTPAHAPGNTATQATOCALT 351
Db 278 PSGOTAPAGSATNVCKAASTST 300

RESULT 9
Q9U013 ID Q9U013 PRELIMINARY; PRT; 560 AA.
AC Q9U013;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Variant-specific surface protein H7-1.
GN H7-1.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS;
RX MEDLINE=96010591; PubMed=7581335;
RA Nash T.E., Conrad J.T., Mowatt M.R.;
RT "Giardia lamblia: identification and characterization of a variant-specific surface protein gene family.";
RL J. Eukaryot. Microbiol. 42:604-609(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GS;
RA Nash T.E., Mowatt M.R., Conrad J.T.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF202776; AAF17594.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR005127; Giardia_VSP.
DR Pfam; PF03302; VSP; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 560 AA; 7023 MW; 6816D55713249BA8 CRC64;

Query Match          9.5%; Score 222; DB 5; Length 560;
Best Local Similarity 25.0%; Pred. No. 6.9e-08;
Matches 118; Conservative 34; Mismatches 156; Indels 164; Gaps 27;

QY 7 LILISLFINELRAVPCDGTQTQAGLTD-----VGAADLGTVCNCRPNFYNGGAAQGE 61
```

```
Db 1 MELLINCLIASTLAGAC---STTQANCAVKECEMVGTEI--CTOCKQNVYPINGVCEAA 55
QY 62 ANGNOPFAANNA---ARGICVPCOIN---RVGSVTNAGDLATLATQCTSTQCTPTGALDD 114
Db 56 ASSNTKCKAANGSDNADQTKCKLLQTFMFKGCCYDKAGE-----TGNLI-- 100
QY 115 GVTDFVRSAAQCCKVKCPNFYNGSGPOGEAPGVQVFAAGAAAGVAAVTSQVPCQLNK 174
Db 101 -CTDAASSTTGCVCATC-----EADG---FFRNPDAAD---TTDSCLIC---- 137
QY 175 NDSPATAGAAQANLATQCSNOCPTGTVDLGGVTLVFNSTATLVCVK-RPNFYNGSGPOGE 233
Db 138 NDT-----TGVTIGESNNAKTYTGITNCVKCKPD-----QLS 170
QY 234 APGVQVFAAGAAAGVAAVTSQVPCQINKNDSPATAGAAQANLATQCT--QCPTGTATQ 291
Db 171 APGTK-----ATCTECASNLTKTDSATP-----ATSCVTAETCKTGYPFN 213
QY 292 DGVTLVFSNSTOCSOCIANFYFNGFNFEAGKSQLCKPVSKTTPAHAPGNTATQATQ--- 348
Db 214 DN-----ADSKRKLAC--NTAANG-----GIDKCAEC--SLTTPASRAGAIIITCKST 260
QY 349 -----CLTTCPRAGT-----VLDDGTSTNFNVASATECKKCSAG 380
Db 261 NSLSPLKDACLTSCPTAGTYETGSPNKKVCTPCHTSCAGCKDDNT-----AASCTACYPG 313
QY 381 FFASKTTGFTAGTDTCTEKTCLTSGATAKVAEATQKVCQ-ASTFEAKFLS 431
Db 314 SVLSYSGDNTKGT-CIAECTGK-----YLENCADGQCTASTAGSKYCS 355

RESULT 10
Q9GQ45 ID Q9GQ45 PRELIMINARY; PRT; 548 AA.
AC Q9GQ45;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Variant-specific surface protein M21-1 (Fragment).
GN M21-1.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AD-1;
RA Mansouri M., Ey P.L.;
RT "A segment of a vsp72-like gene homolog from a type A-I (group 1) Giardia intestinalis isolate.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF298862; AAG37862.1; -.
DR HSP; P00136; 2CV3.
DR InterPro; IPR000564; 2Fe2S ferredoxin.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002350; kazal.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00261; FU; 5.
DR PROSITE; PS00197; 2FE2S-FERREDOXIN; UNKNOWN_1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00282; KAZAL; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 548
SQ SEQUENCE 548 AA; 56557 MW; 578FE4FDA0A2CF0E CRC64;

Query Match          9.4%; Score 219.5; DB 5; Length 548;
Best Local Similarity 24.1%; Pred. No. 1e-07;
Matches 102; Conservative 45; Mismatches 136; Indels 141; Gaps 24;

QY 77 ICVPCQINRVGSVTNAGDLATLATQCTSQ-----CPTGTALDDGVTDFV 120
```

Db	12	LAVTQADKCTGVN	---	TEICTQCRARGVDFGFCVPPFPQAAAGCTEDGVP	--L	65			
Qy	121	DRSAAQCKCKNF	-	YNGGSPQGEA	-	PGVQVFAAGAAAGAAVTSQVPCQLNKN	---	175	
Db	66	DKTAATCGKCGDGL	LLFWGCGYKTESQPGSDI	---	CTAASNGV	-----	CTECN-TKNGLF	116	
Qy	176	DSPATAGAAANLATQCSNQ	-----	CPTGTVLDGVTILVFNTSATLVCVKRPNFY	224				
Db	117	KNPATAPEKRECHLCHD	ATGADGYMGVEGCATCTAPTN	-----	NKGAATCTECQDGY	170			
Qy	225	YNGSGPOGEAPGVQVFAAGAAAGAAVTSQVPCQIN	---	KNDSPATAGAAANLATQCS	281				
Db	171	NDGACKCKVDGC	-	IDCTGA	-----	NOCTCEDGKYLKNQCVDAG	210		
Qy	282	TOCPGTGAIQDGVTLVFSNST	-	QCSOCIANYFFNGFNFAAGKSOCLKCPVSKTTPAHAPG	340				
Db	211	-QCDQGT	-----	YADPTTGCKPC	-----	GITDCATCEYNATISQPOCK	248		
Qy	341	NTATQATQCLTTPAGTV	-	LDGDTSTN	---	FV	---	A	369
Db	249	TCSTSSNNKWKTAADGTT	CTCDDGGCTGNTHFVBSGTNQKLCVPCGDTTNGGVLCNTCS	308					
Qy	370	SATECTKCSAGFFASKTGT	GTAGTDTCTECLTSGATAKVAEATQKVCQASTTFAKF	429					
Db	309	SKTCTKCLDGYDS	-----	GSQVTVCTAC	-----	PGANCATLCE-RYKRCQCTCKPGFF	357		
Qy	430	LSIS	433						
Db	358	LKDS	361						
RESULT 11									
ID	Q9XTJ7	PRELIMINARY;	PRT;	709	AA.				
AC	Q9XTJ7;								
DT	01-NOV-1999	(TremBLrel. 12, Created)							
DT	01-NOV-1999	(TremBLrel. 12, Last sequence update)							
DE	01-JUN-2002	(TremBLrel. 21, Last annotation update)							
DE	Variant-specific surface protein (fragment)								
GN	VSP417-6.								
OS	Giardia lamblia (Giardia intestinalis).								
OC	Eukaryota; Diplomonadida; Hexamitidae; Giardia.								
OX	NCBI_TaxID=5741;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-BRIS-136;								
RA	MEDLINE=99026095; PubMed=9806870;								
EX	Ey P.L., Darby J.M.;								
RT	"Giardia intestinalis: conservation of the variant-specific surface								
RT	protein VSP417-1 (TSA417) and identification of a divergent homologue								
RT	encoded at a duplicated locus in genetic group II isolates."								
RL	Exp. Parasitol. 90:250-261(1998).								
DR	EMBL; U89266; AAD03483.1;								
DR	InterPro; IPR002174; Furin-like.								
DR	InterPro; IPR005127; Giardia_VSP.								
DR	InterPro; IPR001239; Kazal_inhib.								
DR	Pfam; PF03302; VSP; 2.								
DR	PRINTS; PR00290; KAZALINHIB.								
DR	SMART; SM00261; FU; 3.								
FT	NON_TER	709							
SQ	SEQUENCE	709	AA;	71516	MW;	3512BB844B38D134	CRC64;		
Query Match 9.2%; Score 214.5; DB 5; Length 709;									
Best Local Similarity 22.9%; Pred. No. 2.9e-07;									
Matches 129; Conservative 40; Mismatches 182; Indels 213; Gaps 31;									
Qy	14	FINEIRA	-----	VPCPDGTQTQAGLTDVGAADLGT	-----	CVNC	47		
Db	123	FLNPLRANTKDSVVVSCSD	-----	TAGTFTSGKTYRGVOYCERCDDGAVLTDAAAGDAKTRC	178				
Qy	48	RPNFYNGGAAQGEA	-----	NGNQPFANNAARGI	-----	77			

```
QY 96 ATLATQSTOCPTGTALDDGVTDFDRSAAQVKCKPNFYINGSGSQEAPGVQVFAAGA 155
Db 234 KPACTCGGNNYLTAA-DGTTTCVEQSA-----CSPDSFPVENSGQNR---CVLCGDA 284
QY 156 AAAGV-----AAVTSOCVPC-----QLNKNDSPATAGAAQLATQCSNQ-- 194
Db 285 ANGVDVKCAACTPTDQGRIAPTITICTACTNGYKPSADKTTCEAVSNCKTPGCKACSNKG 344
QY 195 -----CPTGTVLDD-----DGVTLVFN-----TSATL-----CVK 218
Db 345 ENEVCTDCDSSTYLTPTSQIDSCAKIGNYYGATEGAKKICKBCTAANCKTCDGQGCQCA 404
QY 219 CRPNFYINGG--SQGEAPGVQVFAAGAAAGAAVTSOCVPCQIINK----- 263
Db 405 CSDGFYKNGDACSPCHE-----SKTCSAGTA---SDCTEPTGKALRYGDDGTGKTC 454
QY 264 -----NDSPATAGAAQLATQSTQCPTGTAT-ODGV-----TLVFSNSTQ- 304
Db 455 GEGCTTCQSGGACKTCGLTIDGASYCS-ECATTTEYPQNGVCAPKASRATPTCNDSPION 513
QY 305 --CSQCIANYF-FNGN-FEA-----GKSOCLKCPVSKTTPAHAPG-----N 341
Db 514 GVGCTCADNYFKNNGCYEIVKYPGKTVCLISAPNGGTCQAADGYKLDGSLTLVCSGCK 573
QY 342 TATQATQCLTTCPAGTVLDDGTSTNFVAS-----ATECTKSAGFFASKTTGFTAGTD 394
Db 574 ECTSSITDC-FTCLDGYVKSASACTKCDSSCCTCNGAATTCACATGY--KT---ASGEG 627
QY 395 TCTECKKLTSGAT 408
Db 628 ACTSCESD-SNGVT 640

RESULT 13
Q24977
ID Q24977 PRELIMINARY; PRT; 1274 AA.
AC Q24977;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cysteine rich protein.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen N., Upcroft P., Upcroft J.;
RT "A Giardia duodenalis gene encoding a protein with multiple repeats of
a toxin homologue.";
RL Parasitology 111:0-0(1995).
DR EMBL; L29079; AAA74587.1; -.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR005127; Giardia_VSP.
DR Pfam; PF03302; VSP; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00261; FU; 16.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 1274 AA; 135720 MW; 321622872A971A32 CRC64;

Query Match 9.1%; Score 212; DB 5; Length 1274;
Best Local Similarity 20.9%; Pred. No. 7.9e-07;
Matches 97; Conservative 42; Mismatches 169; Indels 156; Gaps 21;

QY 23 CPDGTQTQAGLTDVGAADLGTCVNCRPNFYINGSGAAGFANGNQPF---AANNAARGICV 79
Db 851 CVEGNAQQ-----CKTRPGYTIINTDKQCTKDPEAPCNVEGCETCVEGNAQ 897
QY 80 PCQINRVGSVTNAGDLATLATQSTQCPTGTALDDGVTDFDRSAAQVKCKPNFYINGG 139
Db 898 QCKTRPGYTIIN-----TDTKQC-TKDPEAPCNVEGCETCVEGNAQCKTRPGYTIINTD 951
```

```
QY 140 SPQ-----GEAPGVQVFAAGAAAGAAVNTSQCPQLNKNDSPATAGAAQLATQC----- 191
Db 952 TKQCTKDPEAP-----CNTPNCKTCDNPKTDN-----EICTKCDNGD 988
QY 192 ----SNQC--PTGTVLDDGVTLVFENTSATLCVKCRNFYINGSGSQEAPGVQVFAAGAAA 246
Db 989 YLTPINQCPDCTAISG---YYGDTDKCKACNPE-----CA 1022
QY 247 AGVAAVTSOCVPCQIINK-----NDSPATAGAAQLATQCS 281
Db 1023 ECVGPANNOCTACPVGKMLQYTDINTPVNGGTCMDQCSVSTNDGCAECGAIGGTATCS 1082
QY 282 ----TQOPT--GTALQDGVTLVFSNSTQCSQCIANYF-----FNGNFEAGKSQCL-- 326
Db 1083 KCKNTQQAPLNNGCAASSRVAFCATITSGACTKCKNEGFLKDGCGYQTDPRGQKQVCSNA 1142
QY 327 -----KCPVSKTTPAHAPGNTATQATQCLTTCPAGTVLDDGTSTNFVASAETECTKCSAGF 381
Db 1143 OGGNGCKQCTCANGLAASDGNCA-----ECHSTCATGCTAD-----AADCKKTCATGY 1189
QY 382 FASKTTGFTAGTDTCTECTKLTSGATKAVVAEATQKVOCASSTT 425
Db 1190 YKENGDDTAG--LCKKCKSEKI-SGCK-----QCVSSS 1219

RESULT 14
Q24992
ID Q24992 PRELIMINARY; PRT; 557 AA.
AC Q24992;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Variant-specific surface protein H7.
GN VSPH7.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-92244292; PubMed=1574080;
RA Nash T.E., Mowatt M.R.;
RT "Characterization of a Giardia lamblia variant-specific surface
protein (VSP) gene from isolate GS/M and estimation of the VSP gene
reertoire size.";
RL Mol. Biochem. Parasitol. 51:219-227(1992).
DR EMBL; M80480; AAA18202.1; -.
DR HSSP; P19438; INCF.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR005127; Giardia_VSP.
DR Pfam; PF03302; VSP; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 557 AA; 56870 MW; D2898CB41B12A792 CRC64;

Query Match 9.0%; Score 210.5; DB 5; Length 557;
Best Local Similarity 22.9%; Pred. No. 4.4e-07;
Matches 114; Conservative 48; Mismatches 184; Indels 151; Gaps 29;

QY 7 LILISLFINELRAVPCPDGTQTQAGLTD-----VGAADLGTCVNCRPNFYINGGAAQGE 61
Db 1 MFLINCLIASTLAGAC---STTQANCVAEKCEMVGETEI--CTQCKQNVYPINGVCEAA 55
QY 62 ANGN---QPFANNAARGICVPCQ-----INRVGSVTNAGDLATL--ATQSTQCPTGTAL 112
Db 56 ASSNTKCKOKASDEASDQTCGKCLSTFTMYKGGCYDRTGNLGRITCKTEGSDAGKCGACN 115
QY 113 DGVTDVDFRSAA-----QCVKCRPNFYINGSGSQEAPGVQVFAAGAAA-AGVAAVTSQ 166
Db 116 DE--KGFDNPDANNVDSCLSC-----GQATGVTTPGSSTKYKGVAGCAKC 161
QY 167 CVPQLNKNDSPATV---GAQANL-----ATQC--SNOCPGTGVLDDGVTLVFENT 211
```

Db 162 TRPSQISENTGKEATCTECNANLYIKAVSSPTSATSVCVAEDCKTG-----YFPT 212
QY 212 SATL-----CVKCRPNFYNGSPQGEAPGVQVFAAGAAAGVAAVTSQVPCQINKNDS 266
Db 213 TDTTDSKKKCLTC-----STADKGGIDGCSACELLPTTTRASTVLISCSACSTN-NLS 264
QY 267 PATAGAAANLATQCSQCPTGTAIQDGVTLVFSNSTGSCSQCIANYFFNGFNAEAGKSQCL 326
Db 265 P-----LKNFCMODCPAGTYADSNVC---KPHCTSCACKGD-----NTE---SSCT 305
QY 327 KCPVSKTTPAHAPGNATQATQ-----CLTTCPCAGTVLD---DGTSTNFVASATECTKCS 378
Db 306 AC-----YPGSVLSYGTDNKTGCTAEC-TGKYLENCADGGCTATAGSKYCSCK 355
QY 379 AGFF-----ASKTTGFTAGTDTCTCECHK--LTSGA----- 407
Db 356 SGFVPVNGLCVSAETARAAPPSTPDKNG-----VCTACTERYFLESGGCVQAEKPPG 409
QY 408 -TAKVVAEATOKVQAS 423
Db 410 NTLCTTADAGKCTTCAN 426

RESULT 15

Q07317 PRELIMINARY; PRT; 597 AA.
AC Q07317;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Variant-specific surface antigen 1267 precursor (VSP1267).
GN VSP1267.
OS Giardia intestinalis.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_taxid=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 30957 / WB;
RX MEDLINE=92131058; PubMed=1775165;
RA Mowatt M.R., Aggarwal A., Nash T.E.;
RT "Carboxy-terminal sequence conservation among variant-specific surface
RL Mol. Biochem. Parasitol. 49:215-228(1991).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
DR EMBL: M63966; AAA29159.1; .
DR InterPro: IPR000564; 2Fe2S_ferredoxin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR0021174; Furin-like.
DR InterPro: IPR005127; Giardia_VSP.
DR Pfam: PF03302; VSP; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00261; FU; 5.
DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
KW Multigene family; Repeat; Signal; Transmembrane; Glycoprotein;
KW Antigen.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 597 VARIANT-SPECIFIC SURFACE PROTEIN 1267.
FT DOMAIN 18 563 EXTRACELLULAR (BY SIMILARITY).
FT DOMAIN 38 528 23 X 4 AA REPEATS OF C-X-X-C.
FT TRANSMEM 564 591 POTENTIAL.
FT DOMAIN 592 596 CYTOPLASMIC (BY SIMILARITY).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 559 559 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 597 AA; 60646 MW; E107846B8BF72202 CRC64;

Query Match 9.0%; Score 210.5; DB 5; Length 597;
Best Local Similarity 22.6%; Pred. No. 4.7e-07;
Matches 125; Conservative 50; Mismatches 188; Indels 189; Gaps 33;

QY 5 ILLIILISLFINELRAVPCPD-GTQOAGLTD-VGRAADL-----GTC----- 44
Db 4 IAFYLIILSTF-----AVDCKNSGNSCEAGOCDTIGDTEICMQCNQGVKVPINGICTAHSEE 58

QY 45 -----VNCRPN-----FYNGGAAQ-GEANGN-----QPFAANNAAR-- 75
Db 59 AVTNAGCKKNGGTNIEESDKVCGCGNGYFLHKGCGYKIGEAPGNLICADEASNPGARTA 118
QY 76 GICVPCQINRVGSVTNAGDLATATQC-----STOCPT-GTALDDGVTDVFD-----RS 123
Db 119 GVCGAC---KDYKYKNSDAVAT-ADSCIACEDANCATCGGAGENKCTKIDGVFVGATGN 174
QY 124 AAQCVKCK---PNFYNGG-----SPQGEAPGVQVFAAGAAAGVAAVTSQVCP--C 170
Db 175 EGGCIKCDATTGPNYSYKGVAGCAKCEKPKNAGPAKCIETCAADYLTETADEQTSVCVSAVC 234
QY 171 QLNKNDSPAT--AGAQAANLATQCSNQCPGTGTVLDDGV---TLVFNTSAT-----LCVK 218
Db 235 REGKTHPTTDSAGNKKVCVSC-----GTTNNGGIENGCECTSKESAARAGTEITCTK 288
QY 219 CRPNFYNGSPQGEA-----PGVQVFAAGAAAGVAAVTSQVPCQINKNDSFATA--- 270
Db 289 CSS-----NNLSPLGDACLITDCPAGTYAVSGDSGSVCKPCHNTCAGCQTDDRETCTACYP 344
QY 271 -----GAQANLATQCS-----TOCPTGTAIQDGV----- 294
Db 345 GYLLYESNGATGRCKRECTGAFITNCADGQCTANVGSAKYCTOCKDGYAPIDGICTAVA 404
QY 295 -----TLVFSNSTQCSQCIANY-----FFNGNFEAGKSQCL-----KCPVSKT-T 334
Db 405 AAGRDVSVCTATGGKCTACTGNYALLSGGCYNTQTLPKSVCKAVANSNDGKC---KTCA 461
QY 335 PAHAPGNATATQATCCLTCTCPAGTVLDDGTSTNFVASATECTKCSAGFFASKTTGTAGTD 394
Db 462 NGQAPDPATNFCPLCDSTCAECSTKND-----ADACTKCPFGYY---KTG-----N 504
QY 395 TCTECTCKKLTS 406
Db 505 KCIKTESSNNG 516

Search completed: February 11, 2003, 19:47:52
Job time : 32.6571 secs